

The D.W.K. Yeung Condition for Cooperative Differential Games with Nontransferable Payoffs

Anna V. Belitskaia

*St.Petersburg University,
Faculty of Applied Mathematics and Control Processes, Russia, St.Petersburg*

Abstract Irrational behavior proof condition for single player was introduced in Yeung, 2006. The irrational behavior implies the dissolution of cooperative scheme. The condition under which even if irrational behaviors appear later in the game the concerned player would still be performing better under the cooperative scheme was considered in Yeung, 2006. In this paper the differential game with transferable payoffs was considered. In this paper the irrational behavior proof condition for differential games with nontransferable payoffs is proposed.

Keywords: Pareto-Optimality, Time-consistency, Payoff Distribution Procedure.

1. Problem statement

Consider a two-person non-zero-sum differential game $\Gamma(x_0, [t_0, +\infty))$ starting from the initial state $x_0 \in R^2$ at moment $t_0 \in R^1$.

$$\begin{aligned} \rho a_i S_i + \rho b_i = & \frac{1}{8\gamma} \left[\frac{\partial(a_i S_i + b_i)}{\partial S_i} \right]^2 + \pi S_i + \left[\sum_{j \in K_i} (\bar{e}_j \frac{1}{2m_j}) + \frac{1}{2} \bar{e}_i - \right. \\ & \left. - \frac{1}{2\gamma} \sum_{j \in K_i} (a_i \frac{1}{2m_j}) - \frac{1}{4\gamma} \frac{\partial(a_i S_i + b_i)}{\partial S_i} \right] \frac{\partial(a_i S_i + b_i)}{\partial S_i} - \frac{\partial(a_i S_i + b_i)}{\partial S_i} \delta S_i, \end{aligned} \quad (1)$$

The motion equations have the form

$$\begin{aligned} \dot{x} = f_i(x, u_1, u_2), \quad u_i \in U_i \subset R^l, x \in R^2, \quad i = 1, 2 \\ x(t_0) = x^0, \end{aligned}$$

where $u_i \in U_i$ is the control variable of player i and U_i is a compact set.

Let $g_i(x(t), u_1(t), u_2(t))$ be the instantaneous payoff of player $i \in \{1, 2\}$ at moment $t \in [t_0, +\infty)$.

The payoff of player i is given by:

$$\begin{aligned} K_i(x_0, [t_0, +\infty), u_1(\cdot), u_2(\cdot)) = \int_{t_0}^{+\infty} e^{-\rho(\tau-t_0)} g_i(x(\tau), u_1(\tau), u_2(\tau)) d\tau, \\ g_i \geq 0, \quad i = 1, 2, \end{aligned}$$

where ρ is the discount rate. Suppose that payoffs are nontransferable.

If players agree to cooperate, they play an open-loop strategy pair $u^*(t) = (u_1^*(t), u_2^*(t))$ which generates a Pareto-optimal payoff vector in the game

$\Gamma(x_0, [t_0, +\infty))$:

$$\lambda_i(x_0, [t_0, +\infty)) = \int_{t_0}^{+\infty} e^{-\rho(\tau-t_0)} g_i[\tau, x^*(\tau), u_1^*(\tau), u_2^*(\tau)] d\tau, \quad i = 1, 2,$$

where $x^*(\tau)$ is Pareto-optimal trajectory, which has been calculated by substituting the Pareto-optimal open-loop strategy pair $u^*(t) = (u_1^*(t), u_2^*(t))$ in the motion equations (1) and solving of corresponding differential equations.

In the absence of cooperation players stick to the feedback Nash strategies $\bar{u}(t) = (\bar{u}_1(t), \bar{u}_2(t))$. Noncooperative payoff of player i along the noncooperative trajectory $\bar{x}(t)$ for $t \in [t_0, +\infty)$ is defined as:

$$w_i(x_0, [t_0, +\infty)) = \int_{t_0}^{+\infty} e^{-\rho(\tau-t_0)} g_i[\tau, \bar{x}(\tau), \bar{u}_1(\tau), \bar{u}_2(\tau)] d\tau, \quad i = 1, 2.$$

2. Time-consistency

The stringent requirement for solutions of cooperative differential games is time-consistency. Time-consistency of cooperative solution guaranties optimality of chosen optimality principle along the optimal trajectory. Time-consistent solution to nontransferable payoffs requires

- 1) Pareto optimality throughout the game horizon,
- 2) individual rationality throughout the game horizon.

For time-consistency of solution in the game with transferable payoffs the Imputation Distribution Procedure is used.

For time-consistency of solution in the game with nontransferable payoffs the notion of Payoff Distribution Procedure (PDP) was introduced in Petrosyan, 1997:

Definition 1. The function $\beta_i(t)$ is called a Payoff Distribution Procedure (PDP) if it satisfies

$$\lambda_i(x_0, [t_0, +\infty)) = \int_{t_0}^{+\infty} e^{-\rho(t-t_0)} \beta_i(t) dt, \quad i = 1, 2.$$

Definition 2. An admissible Pareto-optimal solution is called time-consistent if there exist a PDP pair $\beta(t) = (\beta_1(t), \beta_2(t))$ such that the condition of individual rationality is satisfied

$$\int_t^{+\infty} e^{-\rho(\tau-t_0)} \beta_i(\tau) d\tau \geq e^{-\rho(t-t_0)} w_i(x^*(t), [t, +\infty)), \quad i = 1, 2,$$

where $w_i(x^*(t), [t, +\infty))$ — is Nash outcome of player $i \in I$ in subgame $\Gamma(x^*(t), [t, +\infty))$. These payoffs are computed with initial condition on the Pareto-optimal trajectory $x^*(t)$.

Theorem 1. *An admissible Pareto-optimal solution is time-consistent if functions $w_i(x^*(t), [t, +\infty))$ are differentiable. There exist a nonnegative function $\eta_i(t)$, $t \in$*

$[t_0, +\infty)$, such that the following conditions are satisfied:

$$\lambda_i(x_0, [t_0, +\infty)) - w_i(x_0, [t_0, +\infty)) = \int_{t_0}^{\infty} e^{-\rho(\tau-t_0)} \eta_i(\tau) d\tau, \quad i = 1, 2.$$

Then PDP $\beta_i(t)$ can be computed by formula (see Petrosyan, 1997):

$$\beta_i(t) = \eta_i(t) + \rho w_i(x^*(t), [t, +\infty)) - \frac{d}{dt} w_i(x^*(t), [t, +\infty)), \quad i = 1, 2. \quad (2)$$

3. Irrational Behavior Proof Condition

Consider the case where the cooperative scheme has proceeded up to time $t \in [t_0, +\infty)$ and some players behave irrationally leading to the dissolution of the scheme. A condition under which even if irrational behaviors appear later in the game the concerned player would still be performing better under the cooperative scheme is the irrational behavior proof condition (Yeung, 2006), which also called the D.W.K. Yeung condition.

The D.W.K. Yeung condition for the differential game with infinite duration $\Gamma(x_0, [t_0, +\infty))$ is described by following inequality:

$$w_i(x_0, [t_0, +\infty)) \leq \int_{t_0}^t e^{-\rho(\tau-t_0)} \beta_i(\tau) d\tau + e^{-\rho(t-t_0)} w_i(x^*(t), [t, +\infty)), \quad i = 1, 2, \quad (3)$$

where $\beta_i(\tau)$ is time-consistent payoff distribution procedure;

x^* is the cooperative optimal trajectory;

$w_i(x_0, [t_0, +\infty))$ is the noncooperative payoff of player i with the initial state x_0 over the time interval $[t_0, +\infty)$;

$w_i(x^*(t), [t, +\infty))$ is the noncooperative payoff of player i with the initial state on the cooperative trajectory $x^*(t)$ over the time interval $[t, +\infty)$.

Differentiating the right hand side of inequality (2) with respect to t leads to the follows sufficient condition of realization of inequality (2).

Theorem 2. *Performance of the following inequality is sufficient to realization of the D.W.K. Yeung condition for player $i = 1, 2$ on time interval $[t_0, +\infty)$:*

$$\beta_i(t) \geq -\frac{d}{dt} w_i(x^*(t), [t, +\infty)) + \rho w_i(x^*(t), [t, +\infty)). \quad (4)$$

Proposition 1. *The D.W.K. Yeung condition for any Pareto-optimal solution in the cooperative differential game with nontransferable payoffs is satisfied for PDP, which is computed by formula (1).*

Proof. If Pareto-optimal solution is time-consistent, PDP $\beta_i(t)$ can be computed by formula (1).

Rewrite formula (1):

$$\begin{aligned} \beta_i(t) - \eta_i(t) &= \rho w_i(x^*(t), [t, +\infty)) - \frac{d}{dt} w_i(x^*(t), [t, +\infty)), \\ \eta_i(t) &\geq 0, \quad i = 1, 2. \end{aligned}$$

Because $\eta_i(t)$ is nonnegative function, the PDP $\beta_i(t)$ satisfies the following inequality:

$$\beta_i(t) \geq \rho w_i(x^*(t), [t, +\infty]) - \frac{d}{dt} w_i(x^*(t), [t, +\infty]), \quad i = 1, 2.$$

Then the sufficient condition of realization of inequality (2) for PDP (1) is satisfied. Hence the D.W.K. Yeung condition is satisfied.

4. Example

As a demonstration the technique of computation the PDP and verification the D.W.K. Yeung condition consider the dynamic game of emission reduction.

The dynamics of the model is proposed in Petrosyan and Zaccour, 2003.

As a Pareto-optimal solution consider the payoffs that implement the maximal total payoff.

4.1. Problem statement

Let $I = \{1, 2\}$ be the set of countries involved in the game of emission reduction.

The game starts at the instant of time $t = 0$ from initial state x_0 .

Denote by $u_i(t)$ the emission of player $i = 1, 2$ at time $t, t \in [0; \infty)$.

Let $x(t)$ denote the stock of accumulated pollution by time t . The evolution of this stock is governed by the following differential equation:

$$\begin{aligned} \dot{x}(t) &= \sum_{i \in I} u_i(t) - \delta x(t), \\ x(t_0) &= x_0, \quad i = 1, 2, \end{aligned} \tag{5}$$

where δ is the natural rate of pollution absorption.

$C_i(u_i)$ is the emission reduction cost incurred by country i while limiting its emission to level u_i :

$$C_i(u_i(t)) = \frac{\gamma}{2}(u_i(t) - \bar{u}_i)^2, \quad 0 \leq u_i(t) \leq \bar{u}_i, \quad \gamma > 0.$$

$D_i(x(t))$ denotes its damage cost:

$$D_i(x) = \pi x(t), \quad \pi > 0.$$

Both functions $C'_i(u_i) < 0$ and $D'_i(x) > 0$ are continuously differentiable and convex.

The payoff function of the player i is defined as

$$K_i(x_0, [0, +\infty), u) = \int_0^{\infty} e^{-\rho t} (C_i(u_i(t)) + D_i(x(t))) dt,$$

where $u = (u_1, u_2)$ is situation in the game; ρ is the common social discount rate.

Each player seeks to minimize its total cost.

Let payoffs be nontransferable in the emission reduction game.

4.2. Solution of the problem

Every player seeks to decrease total costs, so the sign in the irrational behavior proof condition for the emission reduction game will be different.

The D.W.K. Yeung condition for emission reduction game:

$$w_i(x_0, [0, +\infty)) \geq \int_0^t e^{-\rho\tau} \beta_i(\tau) d\tau + e^{-\rho t} w_i(x^*(t), [t, +\infty)).$$

Formulas for a feedback Nash equilibrium, an optimal cooperative trajectory $x^*(t)$, optimal cost of Grand coalition first obtained in Petrosyan and Zaccour, 2003 we present for completeness.

Noncooperative payoff $w_i(x_0, [0, +\infty))$ can be computed with help of system of Hamilton-Jacobi-Bellman equations.

Noncooperative payoff of player i is defined as:

$$w_i(x_0, [0, +\infty)) = \frac{\pi}{\rho(\rho + \delta)} \left(\frac{\pi}{2\gamma(\rho + \delta)} + \sum_{i \in I} \bar{u}_i - \frac{2\pi}{\gamma(\rho + \delta)} + \rho x_0 \right), \quad i = 1, 2. \quad (6)$$

By minimizing the sum of all player's costs, the cooperative payoff of player i (the component of minimal total payoff) is defined as

$$\lambda_i(x_0, [0, +\infty)) = \frac{\pi}{\rho(\rho + \delta)} \left(\sum_{i \in I} \bar{u}_i - \frac{2\pi}{\gamma(\rho + \delta)} + \rho x_0 \right), \quad i = 1, 2.$$

Corresponding cooperative strategies are defined as:

$$u_i^I = \bar{u}_i - \frac{2\pi}{\gamma(\rho + \delta)}. \quad (7)$$

For computation of optimal cooperative trajectory we need to substitute cooperative strategies (6) in motion equations (4).

Optimal cooperative trajectory:

$$x^*(t) = e^{-\delta t} x_0 + \frac{1}{\delta} \sum_{i \in I} u_i^I (1 - e^{-\delta t}).$$

Noncooperative payoff with initial state on optimal cooperative trajectory $x^*(t)$ is defined as:

$$\begin{aligned} w_i(x^*(t), [t, +\infty)) &= -\frac{\pi^2(8\rho + 3\delta)}{2\delta\rho\gamma(\rho + \delta)^2} + \frac{\pi e^{-\delta t} x_0}{\rho + \delta} + \\ &+ \frac{\pi \sum_{i \in I} \bar{u}_i}{\rho\delta} - \frac{\pi \sum_{i \in I} u_i^I e^{-\delta t}}{\delta(\rho + \delta)}. \end{aligned} \quad (8)$$

For computation PDP by formula (1) let compute nonpositive function $\eta_i(t)$:

$$\lambda_i(x_0, [0, +\infty)) - w_i(x_0, [0, +\infty)) = -\frac{\pi^2}{2\rho\gamma(\rho + \delta)^2} = \int_0^{+\infty} e^{-\rho\tau} \eta_i(\tau) d\tau.$$

As a nonpositive function $\eta_i(t)$ we take the function:

$$\eta_i(t) = -\frac{\pi^2}{2\gamma(\rho + \delta)^2}, \quad i = 1, 2. \quad (9)$$

For computation the PDP for emission reduction game substitute the function $\eta_i(t)$ (8) and noncooperative payoff $w_i(x^*(t), [t, +\infty))$ (7) in the equation (1):

$$\beta_i(t) = -\frac{\pi^2(4\rho + 2\delta)}{\delta\gamma(\rho + \delta)^2} + \pi e^{-\delta t} x_0 + \frac{\pi}{\delta} \left(\sum_{i \in I} \bar{u}_i - \sum_{i \in I} u_i^I e^{-\delta t} \right), \quad i = 1, 2. \quad (10)$$

Substitution of time-consistent PDP (9), noncooperative payoff $w_i(x_0, [0, +\infty))$ with initial state t_0 (6) and noncooperative payoff $w_i(x^*(t), [t, +\infty))$ with initial state on the cooperative trajectory $x^*(t)$ (7) implies to the realization of the D.W.K. Yeung condition for emission reduction game.

References

- Petrosjan, L. A. (1997). *The Time-Consistency Problem in Nonlinear Dynamics*. RBCM - J. of Brazilian Soc. of Mechanical Sciences, Vol. XIX, No 2. pp. 291–303.
- Yeung, D.W.K. (2006). *An irrational-behavior-proofness condition in cooperative differential games*. Intern. J. of Game Theory Rew., **8**, 739–744.
- David W. K. Yeung, Leon A. Petrosyan (2006). *Cooperative Stochastic Differential Games*. Springer Science+Business Media, Inc., P. 242.
- Petrosyan, L., Zaccour, G. (2003). *Time-consistent Shapley value allocation of pollution cost reduction*. Journal of Economic Dynamics and Control, **27**, 381–398.
- Petrosjan, L. A. (2003). *Cooperation in differential games*. Conference Book of the XV Italian Meeting on Game Theory and Applications.
- Petrosyan, L. A., Zenkevich, N. A. (2008). *Time-consistency of cooperative solutions in management*. Game theory and applications, Nova Science Publ., NY.
- Haurie, A., Zaccour, G. (1995). *Differential game models of global environment management*. Annals of the International Society of Dynamic Games, pp. 3–24.
- Kaitala, V., M. Pohjola (1995). *Sustainable international agreements on green house warming: a game theory study*. Annals of the International Society of Dynamic Games, pp. 67–88.
- Dockner, E. J., S. Jorgensen N. van Long and G. Sorger (2000). *Differential Games in Economics and Management Science*. Cambridge University Press, pp. 41–85.